R E P O R T R E S U M E S

ED 011 963 24 AA 000 082

THE EFFECTS OF THE EXAMINER AND THE TESTING SITUATION UPON THE PERFORMANCE OF CULTURALLY DEPRIVED CHILDREN. PHASE I--INTELLIGENCE AND LANGUAGE ABILITY TEST SCORES AS A FUNCTION OF THE RACE OF THE EXAMINER. FINAL REPORT.

BY- PHILLIPS, JUDITH
GEORGE PEABODY COLL. FOR TEACHERS, NASHVILLE, TENN

REPORT NUMBER BR-6-1421-DARCEE-VOL-1-NO-3 PUB DATE OCT 66
CONTRACT OEC-6-10-193
EDRS PRICE MF-$0.09 HC-$0.72 18P.

DESCRIPTORS- *TESTING PROBLEMS, *STUDENT TEACHER RELATIONSHIP, *RACIAL DIFFERENCES, SEX DIFFERENCES, *TASK PERFORMANCE, CULTURALLY DISADVANTAGED, INTELLIGENCE TESTS, LANGUAGE TESTS, NEGRO STUDENTS, RACE INFLUENCES, CHILD DEVELOPMENT, *LANGUAGE DEVELOPMENT, POST TESTING, DEMONSTRATION PROJECTS, EARLY EXPERIENCE, NASHVILLE

POST-TEST INTELLIGENCE SCORES OF YOUNG NEGRO CHILDREN OBTAINED FOLLOWING A 6-MONTH LANGUAGE DEVELOPMENT PROJECT WERE COMPARED ACCORDING TO THE RACE AND SEX OF THE EXAMINER ADMINISTERING THE TEST. PRETEST SCORES OBTAINED PRIOR TO LANGUAGE DEVELOPMENT HAD INDICATED THAT NEGRO EXAMINERS PRODUCED SIGNIFICANTLY HIGHER STANFORD-BINET IQ SCORES THAN WHITE EXAMINERS. A SIMILAR, NONSIGNIFICANT DIFFERENCE HAD BEEN FOUND FOR SCORES ON THE ILLINOIS TEST OF PSYCHOLINGUISTIC ABILITIES (ITPA). IN THE PRESENT STUDY, NO DIFFERENCES ASSOCIATED WITH THE RACE AND SEX OF THE EXAMINER AND THE SEX OF THE CHILD WERE OBTAINED WITH 240 STANFORD-BINET POST-TEST SCORES. WHITE EXAMINERS PRODUCED SIGNIFICANTLY HIGHER SCORES THAN NEGRO EXAMINERS WITH 160 ITPA POST-TEST RESULTS. THE GREATEST DIFFERENCE BETWEEN NEGRO AND WHITE EXAMINERS OCCURRED ON THE ITPA VOCAL ENCODING SUBTEST, WHERE WHITES PRODUCED MORE SPONTANEOUS VOCALIZATION. SEX DIFFERENCES WERE MINIMAL AND INCONSISTENT WITHIN THE ITPA SUBTESTS. DIFFERENCES ASSOCIATED WITH THE RACE OF EXAMINER ON PRE- AND POST-TEST SCORES WERE DISCUSSED IN TERMS OF A COMPLEX INTERACTION (KATZ, 1964). THIS INTERACTION BETWEEN RACE OF EXAMINER AND TASK COMPLEXITY CAUSES NEGRO SUBJECTS TO PERFORM BETTER FOR NEGRO EXAMINERS ON COMPLEX TASKS DUE TO THE INTERFERING EFFECTS OF ANXIETY ASSOCIATED WITH WHITE EXAMINERS, EFFECTS WHICH MAY BE REVERSED IF A TASK IS MADE LESS COMPLEX OR THE ANXIETY ASSOCIATED WITH WHITE EXAMINERS IS REDUCED. (JH)

DEMONSTRATION AND RESEARCH CENTER FOR EARLY EDUCATION

A UNIT OF THE

John F. Kennedy Center for Research on Education and Human Development

GEORGE PEABODY COLLEGE FOR TEACHERS / NASHVILLE, TENNESSEE 37203

The Effects of the Examiner and the Testing Situation Upon the Performance of Culturally Deprived Children

Judith Phillips

Contract No. OE-6-10-193 (6-1421)

October 1966

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

Additional support for these investigations was obtained under Grant 5-RII-MH00765 from the National Institute of Mental Health and Grant 66-9174 from the Office of Economic Opportunity.

DARCEE Papers and Reports
Vol. I No. 3 1966

ED011767

FINAL REPORT

The Effects of the Examiner and the Testing Situation Upon the Performance of Culturally Deprived Children. Phase I: Intelligence and Language Ability Test Scores as a Function of the Race of the Examiner

Contract No. OE-6-10-193

Judith Phillips

October 1966

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

Demonstration and Research Center for Early Education

George Peabody College for Teachers

Nashville, Tennessee

Purpose

Recent research on culturally deprived children reveals that their intellectual performance may be affected adversely by racial characteristics of the examiner and the evaluative situation (Forrester and Klaus, 1964).

Katz (1964) has suggested that the race of examiner and level of task difficulty in evaluative situations interact in a manner analogous to the Yerkes-Dodson effect; that is, the optimal intensity of motivation is inversely related to the difficulty of the task. It is assumed that due to the prestige and power associated with them, white examiners arouse more anxiety in Negro subjects than do Negro examiners. Katz and his associates (1964) have shown that male Negro college students perform better with a white examiner on easy performance tasks. When the task is difficult or presented as an intelligence test, Negro subjects do better with Negro examiners. Katz's research further suggests that the race of the examiner affects not only the level of arousal of Negro subjects, but serves as a cue for the tendency to compete or to avoid competition in evaluative situations.

As a part of our programmatic research in the problems of early education, a series of studies has been proposed to ascertain: (1) The manner in which intellectual and school-related performance of culturally deprived Negro children is affected by the race of examiners; (2) the specific intellectual and school-related performances most affected by the race of the examiner; (3) How such differences may be modified through short-term success and failure experiences with examiners of different races; and (4) The manner in which socialization patterns within the culturally deprived Southern Negro family may be related to the differential responsiveness of Negro children to examiners in terms of their race. Dr. Irwin Katz has served as a consultant in the planning of this project.

The initial phase of this research focuses upon the finding that the intelligence test scores of culturally deprived Negro children at the inception of a language development intervention project differed according to the race of the examiner (Dunn and Mueller 1964; Mueller, personal communication). Negro examiners produced significantly higher Stanford Binet and Peabody Picture Vocabulary (PPVT) IQ scores than white examiners ($p < .05$). A similar, but non-significant difference occurred with scores on the Illinois Test of Psycholinguistic Abilities (ITPA). These findings would be predicted by the Katz hypothesis. That is, performance on a complex and difficult task is impaired when the examiner is white rather than Negro.

Six months after the beginning of the language development project post-test measures were taken on the same sample of Negro children. The present study is an analysis of post-test scores on the Binet and ITPA, comparing the relative effectiveness of Negro and white examiners following the short-term intervention experience afforded by the language development project. No comparisons are presented for scores on the PPVT due to the restricted number of examiners administering that instrument.

Method

Examiners: From the total pool of examiners used to test first graders participating in the Cooperative Language Development Project in the Metropolitan Nashville School

System, twelve Negro and white examiners, with equal numbers of men and women, were selected who had administered Stanford Binet post-tests. An additional eight women were selected, half of them Negro and half of them white, who had administered ITPA post-tests. It was not possible to compare sex of examiner effects on ITPA post-tests.

Subjects: From the total number of subjects tested by the above examiners, a random sample of 311 culturally deprived Negro subjects was selected. The sample was equally divided by sex. The mean chronological age of the sample at the time of testing was 7-0 years. Of the total sample, 240 subjects, 120 boys and 120 girls, had post-test IQ scores on the Stanford Binet which were subjected to analysis. The post-test ITPA scores of 160 subjects, 80 boys and 80 girls, were also analyzed.

Procedure: Since this study is based upon data previously gathered as part of the evaluation of the Cooperative Language Development Project, it is essentially a post-hoc analysis.

Stanford Binet post-test IQ scores were subjected to a 2 (Race of Examiner) X 2 (Sex of Examiner) X 2 (Sex of Subject) Factorial Analysis of Variance with 30 subjects in each of the eight cells.

ITPA post-test total standard scores were subjected to a 2 (Race of Examiner) X 2 (Sex of Subject) Factorial Analysis of Variance with 40 subjects in each of the four cells. For the same group of subjects a second analysis of variance was performed on ITPA subtest language age scores according to a Type III (Lindquist, 1956) mixed design: 2 (Race of Examiner) X 2 (Sex of Subject) X 9 (ITPA Subtest). This second analysis enabled a comparison of race of examiner effects between subtests tapping different language abilities (McCarthy and Kirk, 1961).

All major analyses were performed on the IBM 7072 computer. The level of significance for all analyses was $p<.05$.

Results

Analysis of variance of Stanford Binet IQ post-test scores (Table 1) revealed no significant main effects or interactions associated with race of examiner, sex of examiner, or sex of subject. The mean IQ post-test score for the sample was 90.43 (N = 120).

Analysis of variance of ITPA Total Post-Test Standard Scores reported in Table 2 shows a significant main effect due to race of examiner. That is, subjects tested by white females scored higher (mean = -.78) than subjects tested by Negro examiners (mean = -1.44) ($F_{1, 156} = 10.69$, $p<.001$). Neither sex of subject nor the interaction between race of examiner and sex of subject reached significance.

Figure 1 depicts the ITPA profiles of subjects tested by white and Negro female examiners. Table 3 presents the results of the analysis of the nine ITPA Subtest Scores expressed as language age scores. Again, the main effect of race of examiner was significant ($p<.0001$). The main effect of subtest was significant ($F_{8, 1248} = 16.99$, $p<.000001$). There were significant interactions between Subtest and race of examiner

($F_{8, 1248} = 4.61$, $p < .00002$) and subtest and sex of subject ($F_{8, 1248} = 2.26$, $p < .02$). The main effect of sex of subject was not significant.

Table 4 presents the mean language age scores for the significant effect of ITPA Subtest in the analysis of variance. Table 5, presenting the sequential analysis of the mean differences within the significant main effect of subtest (Newman-Keuls procedure, Winer, 1962), reveals that the mean language age score for the Auditory-Vocal Sequencing subtest (80.7) is significantly higher than all other subtest scores ($p < .05, < .01$). Conversely, the mean language age score for the Auditory-Vocal Automatic subtest (63.2) is significantly lower than all other subtest scores ($p < .01$). The mean language age scores on the Vocal Encoding Test (77.0), the Visual Motor Association Test (74.9), the Visual Decoding Test (74.3), and the Auditory-Vocal-Association Test (73.4) are significantly higher than the mean language age score on the Auditory Decoding Test (69.3).

Inspection of means within the significant interaction between ITPA subtest and the race of examiner (Table 6) reflects the tendency for white examiners to produce higher scores on all subtests. The greatest difference between examiners is found in the Vocal Encoding subtest where white examiners produced a mean of 86.4 months and Negro examiners produced a mean of 67.6 months ($p < .01$).

Table 7 indicates that differences between male and female Negro subjects on mean language age scores are not large and contribute equally to the significant interaction between ITPA subtest and sex of subject.

Discussion

The results of this study suggest that a six-month intervention experience in language development has had a modifying effect upon the relative efficacy of Negro and white examiners. The lack of continuing superiority of Negro examiners on the Binet post-test results and the appearance of higher scores produced by white examiners on the ITPA possibly can be attributed to greater familiarity with testing situations and lowered anxiety towards white examiners.

The Katz hypothesis predicts that as stress associated with being tested on a complex task by a white examiner is lowered, an optimal level of motivation is attained and performance level goes up. Under these conditions the level of performance of Negro subjects with Negro examiners is less affected, tending to reduce inter-examiner differences which were initially present due to the tendency of subjects to respond more anxiously to white examiners. It is interesting that the modification of the race of the examiner effect should show up on the ITPA scores, an instrument which would be expected to be most sensitive to changes in performance as a function of a language intervention project.

The higher ITPA scores obtained with white examiners should be considered in the light of the significant interaction between the race of the examiner and ITPA subtest. The greatest difference between white and Negro examiners occurred on the Vocal Encoding subtest, a test of the ability of the subject to express ideas in spoken words. Other studies have found that mentally retarded and culturally deprived children score poorly on this subtest (Gray, Weaver and Starke, 1965; Meuller and Weaver, 1964; and Weaver, 1963).

Because of the post hoc nature of the present study, it is not possible to know if the difference on Vocal Encoding associated with the race of the examiner is due to motivational rather than ability factors in the Negro subjects or to differences in examiner competencies between the two groups of Negro and white examiners.

The present study indicates that although Negro subjects were poorest in performance on the Auditory-Vocal-Automatic subtest, they produced comparatively high scores on the Auditory-Vocal-Sequencing subtest. Part of the reason for this difference may inhere in the complexity of the response required of the subject. The AVSeq subtest requires repetition of digits whereas on the AVAut subtest the subject must complete sentences according to grammatical rules. Negro subjects may be especially handicapped on this test because of racial language patterns that may be at variance with the grammatical constructions of formal English.

Table 7 presents a breakdown of subtest scores according to level of organization, abilities and channels of communication (modes of reception and response). An earlier study by Weaver (1963) reported that culturally deprived children were best on Visual-Motor subtests and poorest on Auditory-Vocal subtests. Comparison of these channels in the present study suggests that this effect occurs in the present sample. Over-all, the subjects appear better at the representational than automatic-sequential level. There is a tendency for all subjects, except those tested by white examiners, to be better in Sequencing than Automatic ability. All subjects, except when tested by Negro examiners, tend to be better on Vocal rather than Auditory channel subtests.

These unanalyzed comparisons serve only to strengthen the suggestion that the results obtained in the present study reflect a complex interaction between the effects of the language program intervention experience, the linguistic patterns of young Negro children, and the effects associated with the race of the examiner. These effects can only be examined by further studies which control the examiner competencies independent of race and in which pre and post tests are done by the same sample of examiners, conditions which did not obtain the data available for the present study. Attention should be given to the effects of the race of the examiner upon the specific linguistic abilities measured by the individual ITPA subtests. At present it may be concluded that the race of the examiner has an effect upon the motivation of Negro subjects and that such effects may operate to obscure the true level of intellectual functioning of Negro subjects in evaluative situations.

Conclusions and Implications

The results of the present study, despite its post hoc nature, support the assumption that the racial characteristics of evaluative adults have important effects upon the performance of young Negro children in task situations. These effects appear to operate in a complex fashion. Of the effects observed, the most significant appears to be that under certain conditions a white examiner may facilitate performance whereas under different conditions the presence of a white examiner may interfere with task performance. Further understanding of these effects obviously will have implications for procedures in the classroom as well as in the structured testing situation.

Summary

Post-test intelligence scores of young Negro children obtained following a six month language development project were compared according to the race of the examiner administering the test.

Pre-test scores obtained prior to the language development project indicated that Negro examiners produced significantly higher Binet IQ scores than white examiners. A similar, non-significant difference was found for scores on the Illinois Test of Psycholinguistic Abilities (ITPA).

In the present study of post-test scores, no differences associated with the race and sex of the examiner and the sex of the child were obtained with the Binet (N = 240). On the ITPA white examiners produced significantly higher total and subtest scores than Negro examiners (N = 160). The greatest difference between Negro and white examiners occurred on the Vocal Encoding subtest; white examiners produced more spontaneous vocalization than Negro examiners. Sex differences were minimal and inconsistent within ITPA subtests. Over-all performance in the present study is consistent with patterns reported in other research on young Negro children using the ITPA.

Differences associated with the race of examiner on pre- and post-test scores are discussed in terms of a complex interaction between race of examiner and task complexity proposed by Katz. That is, on complex tasks, Negro subjects perform less well for white than for Negro examiners due to the interfering effects of anxiety associated with white examiners. These effects may be reversed if the task is made less complex or the anxiety associated with white examiners is reduced.

Further research on the processes assumed to underlie the interaction between race of examiner and task complexity is proposed.

References

Dunn, L.M., & Mueller, M. Cooperative language project. Peabody College, Working Paper #3, October 1, 1964 (mimeo).

Forrester, Bettye J., & Klaus, R.A. The effect of race of examiner on intelligence test scores of Negro kindergarten children. Peabody Papers in Human Development, 1964, 2, No. 7.

Gray, Susan W., Weaver, Ann, & Starke, Christiane. Language abilities of children of the early training project. Informal Report No. 4, Peabody College, 1965.

Katz, I. Review of evidence relating to effects of desegregation on the intellectual performance of Negroes. American Psychologist, 1964, 19, 381-389.

Lindquist, E.F. Design and analysis of experiments in psychology and education. Boston: Houghton Mifflin, 1956.

McCarthy, J.J., & Kirk, S.A. Illinois test of psycholinguistic abilities. Urbana, Illinois: Institute for Research on Exceptional Children, 1961.

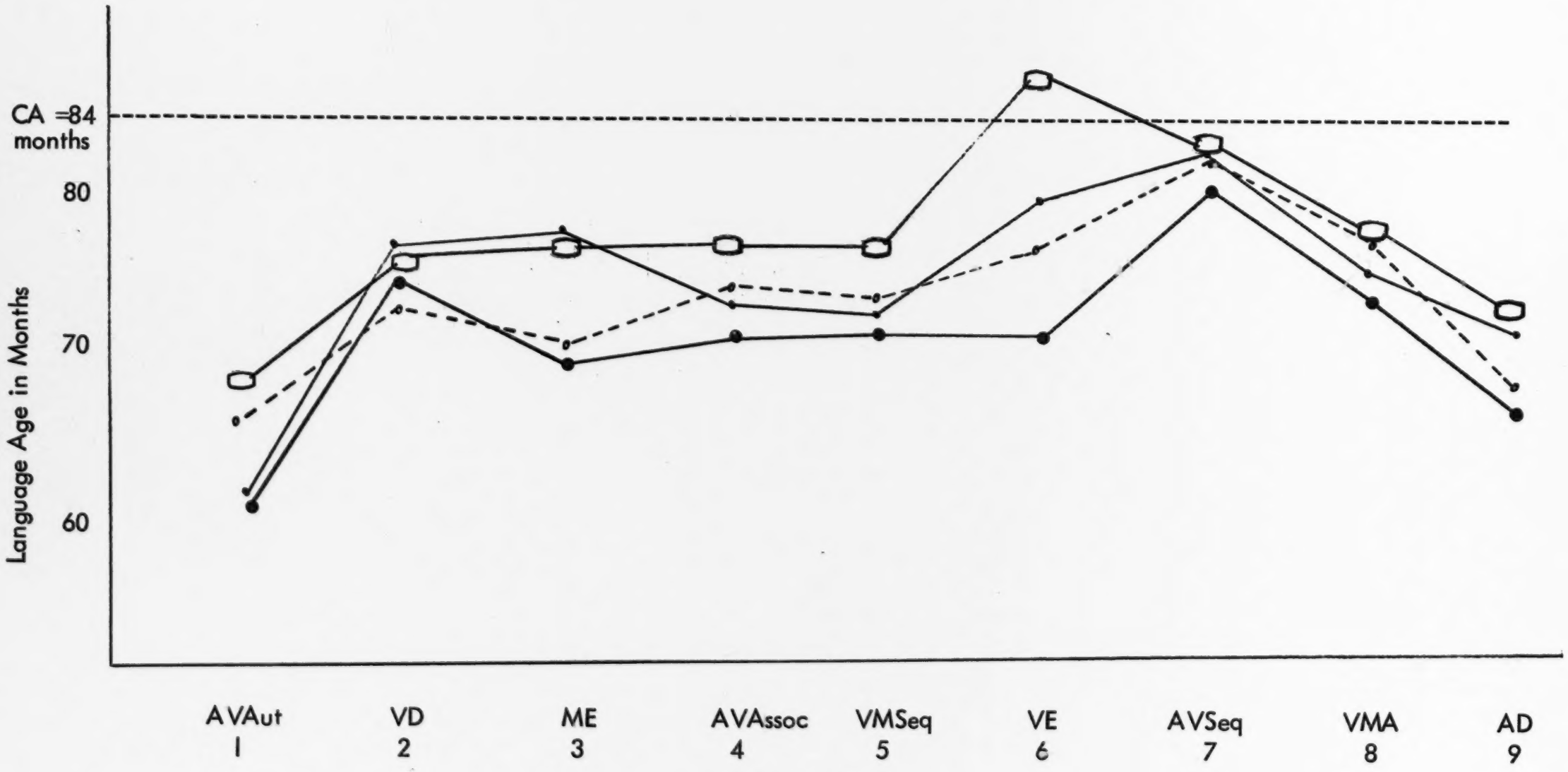
Mueller, M.W., & Weaver, S.J. Psycholinguistic abilities of institutionalized and non-institutionalized trainable mental retardates. American Journal of Mental Deficiency, 1964, 68, 775-783.

Weaver, S.J., Gray, Susan W., & Klaus, R.A. Psycholinguistic abilities of culturally deprived Negro children. Unpublished paper, 1963.

Winer, B.J. Statistical principles in experimental design. New York: McGraw Hill Co., 1962.

Figure 1.

ITPA Post-Test Scores



* All examiners are female, all subjects are Negro

Table 1

Summary of Analysis of Variance of Stanford Binet IQ Post-Test Scores

Source	df	MS	F	p
A (Race E)	1	519.200	2.71	ns
B (Sex E)	1	133.500		
C (Sex S)	1	189.000		
AB	1	451.000	2.35	ns
AC	1	73.700		
BC	1	372.500	1.94	ns
ABC	1	53.300		
ABC/SS	232	191.969		
Total	239	193.845		

Table 2

Summary of Analysis of Variance of ITPA Total Standard Scores

Source	df	MS	F	p
A (Race E)	1	17.44	10.69	.001
B (Sex S)	1	.07		
AB	1	.01		
AB/SS	156	1.63		
Total	159	1.71		

Table 3

Summary of Analysis of Variance of ITPA Language Age Subtest Scores

Source	df	MS	F	p
A (Race E)	1	16865.30	17.47	.0001
B (Sex S)	1	159.20		
AB	1	66.60		
AB/SS	156	965.25		
C (Subtest)	8	3819.50	16.99	.000001
AC	8	1037.34	4.61	.00002
BC	8	507.09	2.26	.02
ABC	8	102.98		
C*AB/SS	1248	224.87		
Total	1439	341.93		

Table 4

Mean Language Age Scores Associated With The Significant Main Effect of ITPA Subtest

	Subtest	Mean Language Age Score
1.	Auditory-Vocal Sequencing (AVSeq)	80.7
2.	Vocal Encoding (VE)	77.0
3.	Visual-Motor Association (VMA)	74.9
4.	Visual Decoding (VD)	74.3
5.	Auditory-Vocal Association (AVAssoc)	73.4
6.	Motor Encoding (ME)	72.6
7.	Visual-Motor Sequencing (VMSeq)	72.3
8.	Auditory Decoding (AD)	69.3
9.	Auditory-Vocal Automatic (AVAut)	63.2
	Mean = 73.1 SD = 5.6 N = 160	

Table 5

Newman-Keuls Sequential Comparison of Differences Between
Language Age Scores Associated With the Significant Main Effect of ITPA Subtest

Source	df	MS	F	p
Subtest (C)	8	3819.50	16.99	.000001
Error	1248	224.87		

Order	1	2	3	4	5	6	7	8	9
Group	AVAseq	VE	VMA	VD	AVAssoc	ME	VMSeq	AD	AVAut
Mean LA Score	80.7	77.0	74.9	74.3	73.4	72.6	72.3	69.3	63.2

	1	2	3	4	5	6	7	8	9
1		3.7*	5.8**	6.4**	7.3**	8.1**	8.4**	11.4**	17.5**
2			2.1	2.7	3.6	4.4	4.7	7.7**	13.8**
3				.6	1.5	2.3	2.6	5.6*	11.7**
4					.9	1.7	2.0	5.0*	11.1**
5						.8	1.1	4.1	10.2**
6							.3	3.3	9.4**
7								3.0	9.1**
8									6.1**
9									

r	2	3	4	5	6	7	8	9
$*q_{95}\ r,\alpha(\sqrt{MSE/n})$	3.32	3.97	4.36	4.63	4.84	5.00	5.15	5.27
$**q_{99}\ r,\alpha(\sqrt{MSE/n})$	4.37	4.94	5.28	5.52	5.71	5.86	5.99	6.10

Table 6

Mean Language Age Scores Within the Significant Interaction
Between ITPA Subtest Scores and Race of Examiner

Subtest	Mean Language Age Scores	
	Negro Examiner	White Examiner
AVSeq	79.9	81.6
VE	67.6	86.4
VMA	72.6	77.2
VD	73.5	75.1
AVAssoc.	70.3	76.5
ME	68.8	76.4
VMSeq.	68.4	76.2
AD	66.4	72.2
AVAut	59.5	67.0
	Mean = 69.7 SD = 5.3 N = 80	Mean = 76.5 SD = 5.1 N = 80

* $p < .05$ (two-tailed t-test)
** $p < .01$ (two-tailed t-test)

Table 7

Mean Language Age Scores Within the Significant Interaction
Between ITPA Subtest Scores and Sex of Subject

Subtest	Mean Language Age Score	
	Male Subject	Female Subject
AVSeq	80.0	81.5
VE	78.4	75.6
VMA	74.0	75.9
VD	75.9	72.7
AVAssoc	72.6	74.2
ME	75.7	69.4
VMSeq	71.3	73.2
AD	71.5	67.0
AVAut	61.3	65.2
	Mean = 73.3 SD = 5.0 N = 80	Mean = 72.7 SD = 4.7 N =80

Table 8

Mean ITPA Language Age Scores by Levels, Abilities and Channels

Measure	Male S	Female S	White Female Examiner	Negro Female Examiner
Levels				
Representational	74.68	72.47	77.30	69.87
Automatic-Sequential	70.87	73.30	74.93	69.27
Abilities				
Decoding	73.7	69.9	73.7	70.0
Association	73.3	75.1	76.9	71.5
Encoding	77.1	72.5	81.4	68.2
Automatic	61.3	65.2	67.0	59.5
Sequencing	77.4	75.7	78.9	74.2
Channels (Single)				
Visual	75.9	72.7	75.1	73.5
Motor	75.7	69.4	76.4	68.8
Vocal	78.4	75.6	86.4	67.6
Auditory	71.5	67.0	72.2	66.4
Channels (Related)				
Visual-Motor	72.8	74.6	76.7	70.5
Auditory-Vocal	71.3	73.6	75.0	69.9